

Three predictions on July 2012 Federal Elections in Mexico based on past regularities

H. Hernández-Saldaña^a

^a*Área de Física Teórica y Materia Condensada, Dpto. de Ciencias Básicas, Universidad Autónoma Metropolitana at Azcapotzalco, Av. San Pablo 180, 02200, México D.F., Mexico.*

Abstract

Electoral systems is subject of study for physicist and mathematicians in last years given place to a new area: sociophysics. Based on previous works of the author on the Mexican electoral processes in the new millennium, he found three characteristics appearing along the 2000 and 2006 preliminary dataset offered by the electoral authorities, named PREP: I) Error distributions are not Gaussian or Lorentzian, they are characterized for power laws at the center and asymmetric lobes at each side. II) The Partido Revolucionario Institucional (PRI) presented a change in the slope of the percentage of votes obtained when it go beyond the 70% of processed certificates; hence it have an improvement at the end of the electoral computation. III) The distribution of votes for the PRI is a smooth function well described by Daisy model distributions of rank r in all the analyzed cases, presidential and congressional elections in 2000, 2003 and 2006. If all these characteristics are proper of the Mexican reality they should appear in the July 2012 process. Here I discuss some arguments on why such a behaviors could appear in the present process.

Key words: vote distribution, election, opinion polls, error analysis, election forensics

1. Justification

Inspired by the courage of Borghesi [1, 2] on made predictions on some behaviors in 2009 elections in France based in past regularities, I present my owns for the July 2012 Mexican election. Unfortunately, the number of datasets onto these predictions are sustained are much more smaller than the French case, but the government's independent institution who is in charge of election organization is relatively new. Additionally, the predictions presented here are made on the Previous Electoral results Program (PREP after its Spanish acronym of

Email address: hhs@correo.azc.uam.mx (H. Hernández-Saldaña)

Programa de Resultados Electorales Previos) which have their present form departing in 2000. On how the PREP works and about its history see the official web page of IFE [3, 4, 5]

This systems offers the first electoral results as the information arrives to the headquarters of authorities. The electoral authorities are grouped in the Instituto Federal Electoral, IFE. The Mexican elections are organized with polling stations distributed along the country and admits, by construction, around 750 votes. The day of the election, the polling stations start at 8 : 00 hrs and close at 18 : 00 hrs except if there are voters remaining. The votes are counted by citizen elected and trained by the IFE in the presence of the political parties representatives. Several data are hand written by the president of the polling station, the highest authority at the cabin, in the cabin certificate. These data are: Total number of received ballots at the beginning of the electoral process (Br), number of remaining (not used) ballots (Bs), number of voters (V), number of deposited ballots per cabin (Bd) and the number of votes received for each party/candidate, V_i . The certified is stamped outside the electoral package, which contain the physical votes. The president carries the electoral package to the collecting stations at the district head, named CEDAT, *Centros de Aco-pio y Transmisión de Datos*. The data are captured and send to the IFE's headquarters.

Hence, according to the IFE's we page[4]: "El PREP NO cuenta votos, sino que captura y publica la información asentada en las Actas de Escrutinio y Cómputo por los ciudadanos que participan como funcionarios de casilla.(The PREP DOES NOT count votes, but it captures and publishes the information seated in the certificates of scrutiny and calculation by the citizens who take part as civil servants of cabin)". And, "Los resultados presentados por el PREP son preliminares, tienen un carácter informativo y no son definitivos, por tanto no tienen efectos jurídicos. (The results presented by the PREP are preliminary, have an informative character and are not definitive, therefore they do not have juridical effects.)"[4]. The definite results are counted during the Count by District, which will start in July 4th. However, PREP offers an opportunity to know and learn about the errors and the ways the democratic process is carried on. In July 2006, the avoided crossing between the two main candidates percentage of votes at PREP's dataset caused the suspicion of a large fraud, mainly by electronic ways. In this sense it is of fundamental importance to understand it and to fulfill the expected reliability. On the process of July 2006 see [6] for a chronicle, and references [7, 8, 9, 11, 10] about the controversial results. Recently appeared a new book [12]. The bibliography is certainly incomplete but I focus mainly in peer reviewed works and sociophysics literature.

Notice that, for the present process, "Por primera ocasión, durante este Proceso Electoral Federal, el IFE pondrá a disposición de la ciudadanía, a través de Internet, la imagen digital de las AEC(Actas de Escrutinio y Cómputo) de las más de 143 mil casillas. (For the first occasion, during this Electoral Federal Process, the IFE will put at the disposal of the citizenship, across Internet, the digital image of the AEC (*voting certificates*) of more than 143 thousand cabins)"[4].

B. received - (B. not used + Number of voters)	Br - (Bs + V)
B. received - (B. not used + B. deposited)	Br - (Bs + Bd)
B. received - (B. not used + Votes for each party)	Br - (Bs + $\sum_i V_i$)
Number of voters - B. deposited	V - Bd
Number of voters - Votes for each party	V - $\sum_i V_i$
B. deposited - Votes for each party	Bd - $\sum_i V_i$

Table 1: Table of errors considered for a self consistency test of the PREP database in July 2000 and 2006. The prediction will run on the same distribution of errors for all the obtainable data during July 2012. We abbreviated Ballots with B.. The variable i stands for the number of votes obtained for each party/candidate.

2. The Predictions

The origin of the author interest in electoral data was due to the suspicion of a Mega fraud in July 2006. But, soon, I discovered the increasing interest of physicist and mathematicians in the field. But there is a lot theoretical work, many of them summarized in the report of Castellano [13], but not so much on real data. Some of the references on actual data are [1, 2, 14, 15, 16, 17, 19, 20, 21, 22, 11]. With the experience gained analyzing electoral data in Mexico, some regularities appeared, and I hope that some of them are part of the Mexican electoral system (I accept that the predictions could be wrong). The main regularities appeared during the studies of Mexican elections [11, 21, 23, 24] are presented in the next subsections.

2.1. 1) Errors could be epidemic in contemporary Mexican elections

Suspicion of a large fraud in July 2006 force us to analyze the self consistency data contained in the dataset provided by IFE. In order to test the existence of anomalies in the presidential data we calculated the error of all the independent tests (summarized in Table 1) and to contrast them with the results of congressional elections of the same year and the presidential one in 2000 [11]. The global behavior, characterized by a power law decay in the center and two asymmetric lobes at each side appears in *all* the cases, with small differences. This result is highly surprising since the presidential process was in complete suspicion of fraud, meanwhile the presidential process in 2000 was clear. Hence I present the following prediction:

Error distributions in self consistency tests of PREP dataset will be described globally by a power law at the center and two asymmetric lobes at each side.

Each distribution of error is builded up by calculating the error defined in 1 on each cabin and doing the histogram of the values obtained, i.e., how many cabins have values of error equal to 0,1,2,... In other words, we are seen for appearance and missing of votes. Notice that in the ideal case all the error distributions must be a Dirac delta function, or no lack or excess of votes.

For the figures see [11] at the arxiv.

2.2. II) The Partido Revolucionario Institucional (PRI) is a sprinter

The PREP data is published in real time and reproduced by several information services. In the dataset the time of arrival to the capture center is recorded, hence a graph of percentage of votes for each party against the time is possible, but it is much more easy to handle if we plot the percentage of processed certificates instead of time. In [11] we found that the behaviour of PRI ruled the general vote decay of the other political parties due to a change in the behaviour. Beyond 70% of processed certificates PRI presents a revival, it changes its rate of grow. This behavior is present in all elections in 2006 (Figure 1 and 2 in reference [11]) and the presidential of 2000 (Figure 9 at the same work). It is a well known fact that PRI receive a lot of votes in geographical regions with a high marginalization index(see for instance [8]), such a regions could have a slow process of cabins to the capture centers, explaining why PRI is a sprinter,i.e., it have a better performance at the end of the journey. During all this years many of such marginal regions are governed by the PRI, hence there is no reason to believe that the mechanism that gives the performance improvement had been missed.

With these arguments I propose the second prediction:

In the graph of percentage of vote against percentage of processed certificates the PRI will change its rate of grow around the 70% of computed certificates. i.e. this political party has a good sprint

2.3. III) The PRI have a smooth vote distribution

During our analysis of electoral data, the smoothness of the vote distribution for the PRI was matter of special atention. This distribution is the histogram of the number of polling stations with certain amount of votes. For comparison with probability distribution the amount of votes is “unfolded” or “deconvoluted” by using the average of number of votes (See reference [23] for an explanation but this procedure is standard in data treatment of complex quantum systems). In [21] I reported the smooth behavior of this party in federal elections 2000, 2003 and 2006 using the definitive dataset of Count by District. A fitting with a model, named daisy[25], of different ranks was tested with success. This model depends only on the rank, r , as free parameter and is written for the nearest neighbour as:

$$P_r(x) = \frac{(r+1)^{r+1}}{\Gamma(r+1)} x^r \exp[-(r+1)x]. \quad (1)$$

With r integer and $\Gamma(\cdot)$ the Gamma function.

However, the distribution of daisy models is a particular case of a more general distribution named gamma and characterized by two free parameters, α and θ [26]:

$$P_\Gamma(x) = \frac{x^\alpha}{\Gamma(\alpha)\theta^\alpha} \exp[-\frac{x}{\theta}]. \quad (2)$$

Here the free parameters are real numbers. Author's hope is that this distribution fits better the distribution of vote, but the daisy models offers a more physical interpretations as explained in [23, 24]. In this way, I present the third prediction:

The distribution of votes for PRI, in presidential and both chambers elections, could be fitted by a smooth distribution, in general by a gamma distribution or for those distributions of Daisy models.

Even when PRI lost presidential election in 2000 and 2004 there were few polling stations with a small number of votes for this political party, hence it is unlike (but not impossible) that such an event occurs in the present election when its presidential candidate had been at the head in all the polls. Notice that even when its candidate shall not win the election the distribution of votes could show a polynomial grow.

3. Acknowledgments and apologies

This work was supported by PROMEP/SEP and projects from CBI-UAM-A. I thank to A. Kunold, J.L. Cardoso and E. Vázquez for encouragement and support. All the errors are my fault. I apologize for the broken English and the lack of appropriate figures. I shall repair it for the next version.

References

- [1] Borghesi, C. *Three quantitative predictions based on past regularities about voter turnout at the French 2009 European election*. arXiv:0905.4578v1 [physics.soc-ph]
- [2] Borghesi, C. *Results of three quantitative predictions based on past regularities about voter turnout at the French 2009 European election* arXiv:0906.2692v1 [physics.soc-ph]
- [3] <http://www.ife.org.mx>
- [4] <http://www.ife.org.mx/informacion-relevante.html>
- [5] All data were obtained from the web page of the IFE, http://www.ife.org.mx/prep2006/bd_prep2006/bd_prep2006.htm
- [6] Klesner, J.L., 2007. The July 2006 presidential and congressional elections in Mexico, *Electoral Studies* **26** 803-808.
- [7] J.A. Crespo. *2006: Hablan las actas. Las debilidades de la autoridad electoral mexicana*. (Debate, 2008). In spanish.
- [8] F. Pliego Carrasco. *El mito del fraude electoral en México*. (Pax México, 2007). In Spanish.

- [9] <http://em.fis.unam.mx/public/mochan/elecciones/indexen.html>.
- [10] J. Aparicio. Nexos. Num. 346, Mexico.(2006)
- [11] G. Báez, H. Hernández-Saldaña and R.A. Méndez-Sánchez. Previous versions of this work. [arXiv:physics/0609144](#) versions 1, 2 and 3.
- [12] Díaz-Polanco. *La cocina del infierno*. México. Temas de Hoy. 2012
- [13] Castellano, Claudio, Fortunato, S., Loreto, V. Statistical physics of social dynamics. *Rev. Mod. Phys.* **81** (2009) 591-646.
- [14] Borghesi C, Galam S. *Phys. Rev. E* **73** (6) (2006) 066118.
- [15] Borghesi C., J-P. Bouchaud. *Eur. Phys. J. B.* **75** (2010) 395-404
- [16] C. Borghesi, J.C. Raynal and J .P. Bouchaud. Election turnout in many countries: similarities, differences, and a diffusive field model for decision making. [arxiv:1201.0524v1](#) Jan 2, 2012
- [17] Christian Borghesi, Jean Chich e, Jean-Pierre Nadal. “Between order and disorder: a ‘weak law’ on recent electoral behavior among urban voters? ” [arXiv:1202.6307v1](#) [physics.soc-ph]
- [18] R.N. Costa Filho *et al* *Phys. Rev E.* **60**, 1067 (1999); *Physica A* **322** (2003) 698.
- [19] A.A. Moreira, D.R. Paula, R.N. Costa Filho and J.S. Andrade Jr., *Phys. Rev E.* **73** (2006) 065101(R).
- [20] S. Fortunato and C. Castellano, *Phys. Rev. Lett.* **99** (2007) 138701.
- [21] H. Hernández-Saldaña. *Physica A* **388** (2009) 2699. [arXiv/physics:0810.0554](#).
- [22] M.C. Gonzalez, A.O. Sousa, and H.J. Herrmann, *Int. J. Mod. Phys. C* **15** (2004) 45.
- [23] H. Hernández-Saldaña. Chapter 16. Traveling Salesman Problem, Theory and Applications. D. Davendra (Ed.) InTech. 2010.
- [24] H. Hernández-Saldaña. M. Gómez Quesada and E. Hernández Zapata. Statistical distributions of quasi-optimal paths in the traveling salesman problem: The role of initial distribution of cities. Accepted in *Rev. Mex. Fís. S.* Available at [arXiv/physics:1003.3913v2](#) [cond.-math.stat-mech] 2011.
- [25] H. Hernández-Saldaña, J. Flores and T.H. Seligman. *Phys. Rev E.* **60** (1999) 449.
- [26] Beyer, W. H. CRC Standard Mathematical Tables, 28th ed. Boca Raton, FL: CRC Press, p. 534, 1987. Weisstein, Eric W. “Gamma Distribution.” From MathWorld—A Wolfram Web Resource. <http://mathworld.wolfram.com/GammaDistribution.html>